

OMNIITOX: LCA Methodology

Evaluation of Selection Methods for Toxicological Impacts in LCA Recommendations for OMNIITOX

Henrik Fred Larsen^{1*}, Morten Birkved¹, Michael Hauschild¹, David W. Pennington² and Jeroen B. Guiné³

¹ Department of Manufacturing Engineering and Management, Technical University of Denmark (DTU) Building 424, DK-2800, Lyngby, Denmark

² Industrial Ecology, Life Cycle Systems, GECOS, Swiss Federal Institute of Technology Lausanne (EPFL), 1015 Lausanne, Switzerland

³ Centre of Environmental Science (CML), Leiden University, P.O. Box 9518, 2300 RA Leiden, The Netherlands

* Corresponding author (hfl@ipl.dtu.dk)

DOI: <http://dx.doi.org/10.1065/lca2004.08.164>

Abstract

Goal, Scope and Background. The aim of this study has been to come up with recommendations on how to develop a selection method (SM) within the method development research of the OMNIITOX project. An SM is a method for prioritization of chemical emissions to be included in a Life Cycle Impact Assessment (LCIA) characterisation, in particular for (eco)toxicological impacts. It is therefore designed for pre-screening to support a characterisation method. The main reason why SMs are needed in the context of LCIA is the high number of chemical emissions that potentially contribute to the impacts on ecosystems and human health. It will often not be feasible to cover all emissions with characterisation factors and, therefore, there exists a need to focus the effort on the most significant chemical emissions in the characterisation step. Until now not all LCA studies include toxicity-related impact categories, and when they do there are typically many gaps. This study covers the only existing methods explicitly designed as SMs (EDIP-selection, Priorfactor and CPM-selection), the dominating Chemical Ranking and Scoring (CRS) method in Europe (EURAM) and in the USA (WMPT) that can be adapted for this purpose, as well as methods presenting novel approaches which could be valuable in the development of improved SMs (CART analysis and Hasse diagram technique).

Methods. The included methods are described. General guidance principles established for CRS systems are applied to SMs and a set of criteria for good performance of SMs is developed. The included methods are finally evaluated against these criteria.

Results and Discussion. Two of the most important performance criteria include providing consistent results relative to the more detailed, associated characterisation methods and the degree of data availability to ensure broader chemical coverage. Applicability to different chemical groups, user friendliness, and transparency are also listed amongst the important criteria. None of the evaluated methods currently fulfil all of the proposed criteria to a degree that excludes the need for development of improved selection methods.

Conclusion and Recommendations. For the development of SMs it is recommended that the general principles for CRS systems are taken into account. Furthermore, special attention should be paid to some specific issues, i.e. the emitted amount should be included, data availability should enable broad chemical coverage, and when identifying priority chemicals for the characterisation, the developed SM should generate few false positives (chemical emissions classified wrongly as being of high concern) and no (significant)

false negatives (classified wrongly as being of low concern) as compared to the associated characterisation method. These recommendations are not only relevant for a stand alone SM, but also valuable when dealing with simple characterisation methods associated with a more detailed characterisation method.

Outlook. There are several questions that need to be answered before an optimal SM can be developed, *inter alia*: Is it optimal to just use simple measured data with high availability or are QSAR estimates of more complex, relevant data better? Which key parameters to include and how? Is a statistical approach, like linear regression of characterisation factors or CART analysis, the best solution?

Keywords: Chemical ranking and scoring (CRS); evaluation criteria; life cycle impact assessment (LCIA); OMNIITOX; selection methods; simple characterisation methods; toxicity-related impact categories

Introduction

Today very few Life Cycle Assessment (LCA) studies include human toxicity and ecotoxicity in the impact assessment, and those which do typically do it in an incomplete way. A reason for this is that in many cases a high number of chemical emissions (termed emissions) potentially contribute to these toxicity-related impact categories ('tox' impact categories), and for most of them there are no available characterisation factors. Applicable tools to deal with this problem do not exist. This calls for methods that are able to select/prioritize those emissions that contribute most significantly to the 'tox' impact categories, where the emissions of consideration are not otherwise limited in the scope of the LCA study. Such methods are here called selection methods (SMs) and their overall aim is to focus the characterisation effort on the most significant chemical emissions when Life Cycle Impact Assessment (LCIA) on toxic releases is performed.

The concept of a selection method has been created within OMNIITOX in order to differentiate from the wider and more ambiguous group of methods covered by the phrase 'LCA screening method'. The latter has been used for concepts with the same purpose as defined for an SM, but the phrase can also mean something quite different, e.g. simple screening level LCA methodology including other impact categories than those which are toxicity related.

The research dealt with here is described in further detail in the OMNIITOX report 'Inventory of LCIA selection methods for assessing toxic releases' (Larsen et al. 2002). The aim of this study has been to come up with preliminary recommendations on development of an SM for use in the classification step of LCIA to support characterisation, i.e. the Base Model (BM) as well as a Simple Base Model (SBM) (see Guinée et al. 2004). It was decided within the OMNIITOX research programme (Molander et al. 2004) to aim at performing the selection with a characterisation method (in this the SBM) which is based on a statistical derivation from the BM using QSAR data to some extent (Guinée et al. 2004). This approach supplements the different approaches to developing SMs which are exemplified in this paper. In addition to its function as a characterisation method, the SBM thus also serves as a selection method, and most of the results presented for the SM are equally applicable to simple characterisation methods like the SBM: the evaluation criteria developed, the recommendations on method development, the analysis of Chemical Ranking and Scoring (CRS) methods and general CRS principles in an LCIA framework. Regardless whether the selection is performed using a simple characterisation method like the SBM or an individual SM, the study described here is an indispensable part of the basis for the development of the selection approach.

In this paper we first describe the characteristics of an SM, followed by a description of existing SMs and other relevant CRS methods. General principles for CRS methods are applied to the special demands on an SM. Criteria for the evaluation of SMs are then presented, and afterwards the included methods are evaluated. Properties of the different methods in the context of the SM approach are then discussed.

1 Selection Methods

An SM can be considered as a CRS method with some special constraints, the most important being the requirement for consistency with the associated characterisation method for which it is supposed to identify the most important emissions (consistency between the lower and higher levels of complexity in the assessment). The aim of the SM is to select (prioritize) those emissions from an inventory of a specific LCA-study that are most likely to contribute – according to a specific characterisation method – significantly to the impact categories on ecotoxicity and human toxicity. In this way, the SM serves for the step of the Life Cycle Impact Assessment (LCIA) called classification or in ISO 14042 terminology: Assignment of emissions to impact categories.

Let's say that we were able to perform the characterisation (i.e. calculate impact potentials) for all the emissions of chemicals mapped in an inventory, which could be several hundred chemicals of interest in some studies, and then subsequently rank their impact potentials. The ideal SM would rank the emissions in the same way (or even better, give the same relative value to each emission) as the characterisation method, but based on a significantly lower data demand.

More pragmatically, the SM does not need to rank the emissions in exactly the same way as the characterisation method,

as long as the significant contributing emissions are subsequently selected for inclusion in the characterisation step. Imperfection in ranking may occur as false positives, i.e. emissions that are identified for further assessment as high concerns, ultimately result in low concern through further evaluation. This is somewhat undesirable because it reduces the ability of the SM to limit the characterisation work. The opposite of false positives, false negatives, are even less desirable because emissions of high concern will not be prioritized for characterisation. As SMs must be based on limited information, a trade-off is often sought by evaluating how many false positives relative to how many false negatives are acceptable, i.e. where to draw cut-offs between high and low potential for concern.

Another important aspect of an SM is the data demand, and thereby the time demand, which preferably should be significantly less than for the full characterisation method – otherwise one could just as well skip the selection and go directly to characterisation. As is the case with characterisation methods, SMs should be applicable to all types of chemicals (in principle any chemical emitted and mapped in the inventory). Furthermore, for each chemical, the substance data required for the use of the SM should be available for all emissions encountered in the inventory. Exceptions may include, for example, metals or other elements for which simplistic SMs are not always applicable and where more detailed measures like characterisation factors may already exist.

2 Existing SMs and other CRS Methods

Since very few SMs actually exist with a direct link to more detailed/robust characterisation methods, it was decided to expand the scope of the analysis to comprise other CRS methods developed for a similar screening purpose within environmental risk or hazard assessment, i.e. prioritizing/ranking of emissions and/or chemicals. It was anticipated that these methods could be adapted, or could contribute with applicable elements and principles, to the development of an SM and the development of SBM in the OMNIITOX research programme.

The number of CRS methods used within hazard assessment and risk assessment is very high. According to Davis et al. (1997) hundreds of methods exist (or have existed) which can be based on different procedures: pre-screening (scientific judgement), hazard ranking approach (threshold effects, no exposure included), ordinal assignment of data (scores, algorithm that weights) and risk-based quotient (including both exposure and effects, and a margin of safety). The basic principles in most of the recent CRS methods include some kind of assessment of the persistence, bioaccumulation and (eco)toxicity of the chemicals in question (so-called PBT approaches).

Table 1 presents a list of chosen CRS methods, including all known SMs and one very simple characterisation method (CPM). The chosen CRS methods used outside LCIA are expected to represent the most important methods, or types of methods, representing the situation in the European Union (EURAM) and in the USA (WMPT). Two mathematical/statistical tools (Hasse and CART), until now only used

Table 1: List of existing SMs and chosen CRS- and mathematical/statistical methods

Method name	Includes	Typical use	Reference
EDIP-selection method	Ecotoxicity and human toxicity	LCIA	(Hauschild and Wenzel 1998, Wenzel et al. 1997)
Priofactor	Ecotoxicity and human toxicity	LCIA	(Larsen et al. 1999b and 1999a)
CPM-method; qualitative part	Ecotoxicity	LCIA	(Eriksson 1999)
CPM-method; quantitative part ^a	Ecotoxicity	LCIA	(Eriksson 1999)
EURAM	Ecotoxicity and human toxicity	Risk Assessment	(Hansen et al. 1999)
WMPT (adopted also in P2P)	Ecotoxicity and human toxicity	Hazard Assessment (and LCIA)	(US EPA 1997, Pennington et al. 2002)
Hasse diagram technique	Ecotoxicity and human toxicity ^b	(Risk Assessment)	(Halfon et al. 1996, Sørensen et al. 1998)
CART	Ecotoxicity and human toxicity ^b	(Risk assessment)	(Bennett et al. 1999, 2000 and 2001)

^a Simple characterisation method^b Can be used for both

to a limited degree within risk assessment, are also included because they are considered to represent promising alternative approaches for the selection of chemicals. In the table, the typical use is given, and it is indicated whether the method includes the evaluation of ecotoxic effects, of human toxic effects, or both.

A method that is not shown in Table 1 but which may become relevant for the development of an SM and/or the planned development of a SBM within the OMNIITOX project (Guinée et al. 2004) is the 'log linear regression equation' method based on a multivariate analysis of the associated characterisation method. A preliminary version of the method was presented by van de Meent et al. (2002), but sufficient documentation was not available in order to prioritize the method for further evaluation in this study. Furthermore, CRS methods based on risk phrases for human toxicity to be used in connection with LCIA have recently been published (Bunke et al. 2003). However, these methods are not included here because the use of risk phrases for ranking was already adopted in the EDIP-selection method, the Prio-factor, and EURAM.

Each method is shortly described below. A more detailed description of all the methods can be found in Larsen et al. (2002) or in the literature related to the method.

2.1 EDIP-selection method

The EDIP-selection method was developed during the creation of the EDIP method (Wenzel et al. 1997). This is a full LCA method including tools, for example, for calculating key property based characterisation factors for ecotoxicity

Table 2: Exposure scores for the EDIP-selection method

Emission expected	Classified as R53 or R58	Score
Yes	Yes	8
Yes	No	4
No	Yes	4
No	No	1

R53: May cause long-term adverse effects in the aquatic environment

R58: May cause long-term adverse effects in the environment

and human toxicity. In the EDIP-documentation, this SM is described as a qualitative/semi-quantitative tool for screening chemical emissions and for prioritization of chemicals to be included in the quantitative characterisation step of the EDIP method.

The EDIP-selection method is based on risk phrases, i.e. the R-sentences used in EU labelling of chemicals based on hazard assessment (ECC 1967 and its amendments, e.g. EC 2001) and an assessment of whether the chemical is emitted or not. For each chemical, two impact scores are calculated, one for ecotoxicity and one for human toxicity, by multiplying an exposure score with an (eco)toxicity score.

The ecotox impact scores (Table 4) are calculated by multiplying an exposure score (Table 2) with an ecotoxicity score (Table 3). The R53 and R58 used in the criteria for the exposure score are both based on the persistence and/or bioaccumulation potential of the substance in question. The ecotoxicity score is, in practice, only based on acute aquatic ecotoxicity (R50, R51, R52) because assignment of terrestrial risk phrases are typically missing.

Table 3: Ecotoxicity scores for the EDIP-selection method (final ecotoxicity score equals the sum of the two scores, i.e. 8, 6, 5, 4, 2, 1, 0)

Ecotoxicity	Aquatic ecotoxicity				Terrestrial ecotoxicity	
	Criteria	R50 or R50/53	R51/53	R52 or R52/53	N.C.	R54 or R55 or R56 or R57
Score	4	2	1	0	4	0
R50: Very toxic to aquatic organisms					R55: Toxic to fauna	
R51: Toxic to aquatic organism					R56: Toxic to soil organisms	
R52: Harmful to aquatic organisms					R57: Toxic to bees	
R53: See notes in Table 2					N.C.: No Classification	
R54: Toxic to flora						

Table 4: Ecotox impact scores for the EDIP-selection method

Exposure score	Ecotoxicity score						
	0	1	2	4	5	6	8
1	0	1	2	4	5	6	8
4	0	4	8	16	20	24	32
8	0	8	16	32	40	48	64

For the calculation of the human impact scores the same exposure score as for ecotox is used (see Table 2). The human toxicity part is also based on risk phrases (e.g. R20-28, R40, R45, R46, and R49), but here several effect categories are included, e.g. acute toxicity, genotoxicity and carcinogenicity (Larsen et al. 2002).

As is evident from the description, the EDIP-selection method can be characterised as a PBT approach based on regulatory hazard assessment and only including emissions as yes or no (not the emitted amount).

The EDIP-selection method has only been evaluated against the EDIP characterisation method in one case study on a detergent for manual dishwashing (Larsen et al. 1999a). The study included 42 emissions. The ecotoxicity part of the EDIP-selection method was not able to discriminate amongst all emissions and resulted in 29 false positives. The main reason for this was the lack of data for terrestrial ecotoxicity, which has a conservative effect on the scoring, i.e. the highest score for terrestrial ecotoxicity (4) was given, see Table 3. For the human toxicity part of the EDIP-selection method, 33 emissions were prioritized resulting in 10 false positives and 5 false negatives.

2.2 Priofactor

The Priofactor is an SM developed for the EDIP method that was developed due to the poor ability of the EDIP-selection method to prioritize emissions, as described in section 2.1. Instead of optimising the risk phrase approach, as used in the EDIP-selection method, it was decided to use a key parameter approach much closer in principle to the EDIP method. The Priofactor is therefore an example of an SM that can be used as a characterisation method, but with a lower demand on data quality and the amount of data than the associated 'higher tier' characterisation method, i.e. the EDIP method. The development of the Priofactor is described in Larsen et al. (1999a) and Larsen et al. (1999b).

The Priofactor is divided in three impact potential categories and therefore involves three calculations per emission – one for potential ecotoxicological contribution and two for potential human-toxicological contribution (direct via air and indirect via water/soil). These calculations require readily available data on (bio)degradation, bioaccumulation, (eco)toxicity, and an estimate on the emitted amount to predict the relative contribution from each substance to the total toxicological impact potential for each category. The Priofactor makes it possible to select those emissions that

contribute with more than, for example, 0.1% or 1% to the total impact potential for each category. In this way, the more time-consuming detailed impact assessment in the characterisation step can be restricted to those emissions.

The overall principles for calculating the Priofactor are:

$$\text{Priofactor} = Q \cdot \text{Subprio} \quad (1)$$

Q is the emitted amount and a subprio is estimated for each of the three categories i.e. ecotoxicity, human toxicity air and human toxicity water/soil. The ecotoxicity subprio (m³/g) is calculated as:

$$\text{Ecotox - subprio} = \frac{\text{BDF} \cdot \text{BCF}}{\text{EC}_{50}(\text{g/m}^3)} \quad (2)$$

Where BDF is the BioDegradationFactor (values between 0.1 and 1 depending on persistence), BCF is the BioConcentrationFactor and EC₅₀ is the concentration of the substance in question having an acute (lethal) effect on 50% of test organisms in an OECD standard acute laboratory test.

The human toxicity subprios are calculated on almost the same principles, but instead of EC₅₀s, Toxfactors are used. The Toxfactors can have values between 2 and 2000 (air) or 200 and 2000 (soil/water) depending on the severity of the classification (i.e. assigned risk phrases) of the substance in question. For the human toxicity subprio (air), a factor between 0.1 and 1.0 based on the half-life in air of the substance is used instead of BDF.

On the basis of the description, the Priofactor can be characterized as a PBT approach that takes the emitted amount into account in a quantitative way.

The Priofactor was evaluated against the EDIP method in the same case study as the EDIP-selection method (Larsen al 1999a). For the ecotoxicity part, the Priofactor selection resulted in four false positives and three false negatives, but importantly here is that the chemical emissions prioritized by the Ecotox-priofactor accounted for 99.63%, whereas the three false negatives only accounted for 0.31% of the total impact potential of the EDIP characterisation. For the human toxicity part the results of the evaluation was almost as good as for the Ecotox-priofactor (Table 5).

Table 5: Results of evaluation of the Priofactor against the EDIP method

	Ecotox-priofactor		Human-tox-priofactor (air)		Human-tox-priofactor (water/soil)	
	Number	Relative share of total EDIP impact potential (%)	Number	Relative share of total EDIP impact potential (%)	Number	Relative share of total EDIP impact potential (%)
False positive	4	0.00	1	0.00	6	0.00
False negative	3	0.31	13	3.15	3	1.23
Prioritized emissions	14	99.63	9	96.77	15	98.74
Total emissions	42	100	28	100	42	100

2.3 CPM

The CPM method is described in Eriksson (1999). It consists of a first screening part (i.e. the qualitative part) that could be regarded as an SM (hereafter called CPM-selection method). In the second part of the method (i.e. the quantitative part termed the CPM-characterisation method) characterisation factors for chronic ecotoxicity in water are calculated, but the method is not included in a full LCIA method.

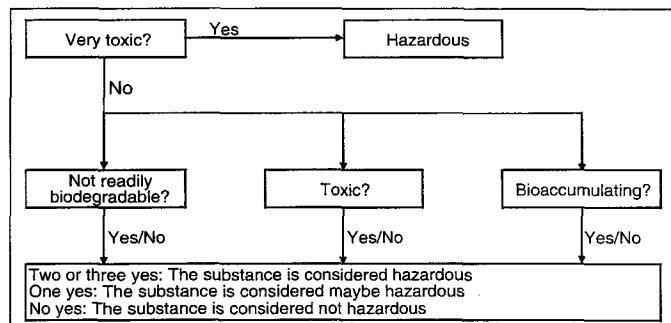
The description of the CPM-characterisation method is only included here to illustrate how a close connection between the data needed for the selection and the characterisation can be achieved and to show a special approach on how to deal with inorganics. As the CPM-characterisation method is a characterisation method, it is not included in the evaluation of SMs in section 5.

Both the CPM-selection method and the CPM-characterisation method are divided into a decision tree/algorithm that handles organic substances and another decision tree/algorithm that handles inorganic substances (mainly metals). The individual methods are described in the following subsections.

2.3.1 CPM-selection method

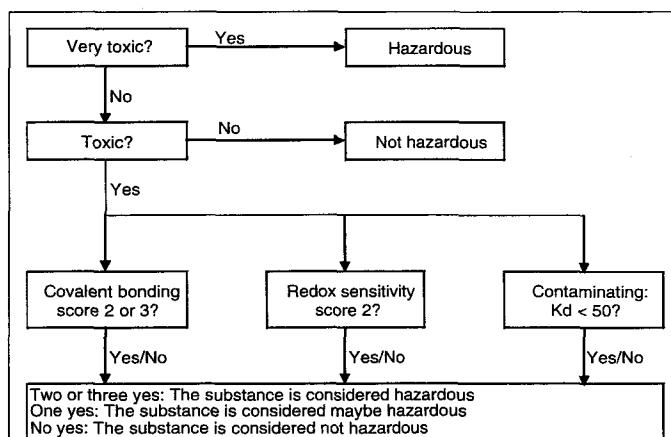
The CPM-selection method on organic substances makes use of a decision tree based on the three intrinsic substance properties ecotoxicity, biodegradation and bioaccumulation (Fig. 1). The criteria for the thresholds (i.e. question marks in Fig. 1) are identical to the ones used in the EU labelling of chemicals based on hazard assessment (EEC 1967, and its amendments, e.g. EC 2001), see Table 6 showing criteria for ecotoxicity. Only substances considered as not hazardous are not prioritized (see Fig. 1).

The decision tree for inorganics is shown in Fig. 2. For this group of substances, biodegradation and bioaccumulation is excluded and substituted by covalent bonding properties, redox sensitivity and properties of partitioning between water and solid phase (K_d). For covalent bonding, soft acids (strong forming, i.e. covalent bounds with ligands) are given the score 3, whereas hard acids (weak forming, i.e. electrostatic bounds with ligands) are given the score 1. Substances in between, i.e. borderline acids, are given the score 2. For redox sensitivity, substances that can change their valence state is given the score

**Fig. 1:** Decision tree for the selection of organic substances by the CPM-selection method

2, others are given the score 1. Substances with a K_d value below 50 are considered as contaminating (i.e. bio-available), otherwise not contaminating. Only substances considered as not hazardous are not prioritized (see Fig. 2).

The CPM-selection method can be characterised as a very simplistic PBT approach, not taking emitted amount into account, but paying attention to some of the special properties of inorganics. An evaluation of this SM against the CPM-characterisation method has, as far as known to the authors, never been published.

**Fig. 2:** Decision tree for the selection of inorganic substances by the CPM-selection method**Table 6:** Ecotox criteria for the CPM-selection method

Test organism	Test result	CPM evaluation	Test result	CPM evaluation
Fish, LC50 (mg/l)	< 1.0	Very toxic	< 100	Toxic
Daphnia, EC50 (mg/l)	< 1.0	Very toxic	< 100	Toxic
Algae, EC50 (mg/l)	< 1.0	Very toxic	< 100	Toxic

2.3.2 CPM-characterisation method

The CPM-characterisation method is a quantitative method that calculates a characterisation factor (called impact factor) for each substance based on its intrinsic properties. It consists of two parts, one for organic substances and one for inorganic substances. For the organic part, the impact factor is calculated in the following way:

$$\text{Impact factor (organics)} = \frac{\text{Log} K_{\text{ow}} \cdot H}{\text{NOEC} \cdot \text{DEG}} \quad (3)$$

Where K_{ow} is the octanol/water partitioning coefficient, H is the Henry's law constant, NOEC is the chronic No Observed Effect Concentration and DEG is the percentage biodegraded in an OECD test for ready biodegradability. H is only included in the calculation if the atmospheric half-life of the substance is over 80 days. For the inorganic part, the calculation is performed in the following way:

$$\text{Impact factor (inorganics)} = \frac{\text{CB} \cdot \text{RS}}{\text{NOEC} \cdot \log K_d} \quad (4)$$

Here, CB is the covalent bonding score (1, 2 or 3), RS is the redox score (1 or 2) and K_d the partitioning coefficient, all three defined in the previous section.

2.4 EURAM

The EU Risk Ranking Method (EURAM) (Hansen et al. 1999) was developed within the EU as a part of the second of four steps, i.e. priority setting, in the evaluation of existing chemicals: Data collection, priority setting, risk assessment and risk reduction. It is used for screening chemicals for selection of the potentially most problematic ones for subsequent risk assessment. The target of the method has been the 2465 high production volume chemicals (HPVC's) for which data are compiled in the European Conform Chemical Information Database (IUCLID 2000). EURAM prioritizes chemicals on the basis of their potential risk to humans (human toxicity) and the environment (ecotoxicity). It is to a large extent consistent with the EU risk assessment methodology described in the Technical Guidance Document (TGD) (EC 1996).

The environmental fate or distribution part of EURAM includes a multimedia equilibrium partitioning model and calculates a score for the Predicted Environmental Concentration (PEC), whereas the environmental effect scoring is based on PNEC-values (Predicted No Effect Concentration).

The human effect scoring is based on R-phrases, and the fate part is simple and based on discrete exposure fraction values depending on the boiling point, vapour pressure and K_{ow} of the substance in question. Although EURAM includes methodologies for other protection goals than humans and the aquatic environment, i.e. soil, top predators, the atmosphere and microorganisms in sewage treatment plants, these are not included in this paper on SMs. This is because effect

data for these protection goals are very limited and the EU therefore only uses these scores for prioritization in conjunction with expert judgement.

As far as the authors are aware, EURAM has not been used directly as an SM in LCIA. Though more advanced in some parts, it is, however, based on some of the same basic principles as the previously described SMs and may therefore be relevant as a potential basis for an LCIA selection tool. For aquatic ecotoxicity, an Aquatic Score (AS) is calculated in the following way:

$$\text{AS} = [1.37(\log(\text{Emission} \cdot \text{Dist}_{\text{env}} \cdot \text{Deg}) + 1.301)] \cdot [0.7(-2 \cdot \log \text{PNEC}) + \text{AP}] \quad (5)$$

Where Emission is the emitted amount (here truncated at min 50 ton and max 1,000,000 ton), Dist_{env} is the fraction distributed to the aquatic environment (here truncated at min 0.01 in an EURAM unit world), Deg is a score for biodegradation (0.1; 0.5 or 1.0 depending on biodegradation in OECD tests on readily and inherent biodegradability), PNEC is the Predicted No Observed Effect Concentration and AP is the Accumulation Potential (values of 0; 1; 2 or 3 depending on the BCF value of the substance). The constants within formula 5 are included to ensure that AS gets values between 0 and 100.

The EURAM aquatic score can be characterized as a PBT approach taking both emitted amount and distribution, i.e. the amount ending up in the aquatic environment, into account.

For human health, a Human Health score (HS) is calculated in the following way:

$$\text{HS} = [1.785(\log(\text{Emission} \cdot \text{Dist}_{\text{HH}}) + 0.398)] \cdot \text{HEF} \quad (6)$$

Where Dist_{HH} is a fraction between 0.05 and 1 based on threshold values for the boiling point, vapour pressure and $\log K_{\text{ow}}$ of the substance and HEF is the Human health Effect score assigned values between 0 and 10 depending mainly on the risk phrases assigned in combination with results from some specific laboratory tests, for example, genotoxicity and reproductive toxicity. The constants included in formula 6 are there to truncate the HS between 0 and 100.

The EURAM human health score can be characterized as a BT approach (persistency (P) not included) based on regulatory hazard assessment for human health, but including emitted amount and a threshold based factor for the fraction contributing to human exposure.

2.5 Waste minimization prioritization tool (WMPT)

In 1997, the US EPA's Office of Solid Waste (OSW) and EPA's Office of Pollution Prevention and Toxics (OPPT) issued a beta test version (1.0) of the WMPT (US EPA 1997, 1998). This methodology provides a screening-level assessment of the potential chronic hazard of chemicals in the context of human health and the aquatic environment. This is typical of the approaches adopted by many other national and international organisations. Previous tools also accounted for emission quantity, but this was not explicitly retained in the 1997 WMPT.

The main purpose of the WMPT system was to screen across the thousands of chemicals that may potentially appear in hazardous waste streams in terms of their Persistence (P), Bioaccumulation (B), and Toxicity (T). To conduct this broad-based PBT screening, the WMPT makes use of data from a variety of EPA and outside sources following data source preference hierarchies. For example, Reference Dose (RfD) and Reference Concentration (RfC) data from the Agency's Integrated Risk Information System (IRIS) were considered to be highest preference for toxicity (T) scoring for human non-cancer effects. If these data were unavailable, the tool uses data from lower-preference sources, such as acute toxicity LC₅₀ (lethal concentration killing 50%) data for rodents from non-peer reviewed databases.

Data from various sources are placed on similar scales by comparing data values with established fencelines (or cut-offs) to place the values into low, medium, and high concern bins (or categories) for each of P, B, and T. For example, if a chemical has a bioconcentration factor (BCF) greater than, or equal to, 1000 then it is placed in the high-potential bin and assigned a bioaccumulation (B) score of 3. Similarly, values from 250 to 999 were placed in the medium bin (a medium score of 2 indicating a possibility for both high or low concern), and values below 250 were placed in the low concern bin (score of 1).

As far as possible, the data fencelines between categories were made consistent with the EPA's protocols at the time for screening new and existing chemicals under the Toxic Substances Control Act (TSCA). Where Agency precedents were unavailable for establishing fencelines, they were set so that 25% of the values would be considered high potential for concern, 50% medium, and 25% low (1:2:1). We note that there are a number of pros and cons to this type of simple binning approach, but, on balance, the EPA felt that a three-bin system was the best approach to place disparate types of data on comparable scales and that this was generally consistent with the degree of uncertainty associated with the data.

Comparisons with mechanistic-model based characterisation approaches were presented in Pennington and Bare (2001) and, in the context of selection/classification in LCA, in Pennington and Bare (2003). The WMPT method can be characterized as a PBT approach not taking the emitted amount into account. However, a way to account for the emitted amount at each level of PBT concern was demonstrated (Pennington and Bare 2001). This enabled chemicals of similar concern to be grouped and only those emitted in higher quantities in each group which are to be considered further. Some low concern groups could additionally be eliminated, depending on the scope of the LCA study.

2.6 Hasse diagram technique

In contrast to most other screening methods used, the Hasse diagram technique (HDT) is a purely objective screening method based on a mathematical method called partial order. The HDT has been used in several ways to screen and

select chemicals of concern (Sørensen et al. 1998, Formsgaard and Sørensen 1999, Lerche et al. 2002) and for ranking of sediments (Brüggemann et al. 2001). The use of the method was tremendously facilitated by the development of the WHASSE programme (Brüggemann 1999).

As an example, two chemicals (1 and 2) can be compared in relation to potential environmental risk using the parameters K_{ow}, DT50 (half-life) and 1/(EC₅₀). If chemical 1 possesses a potentially higher danger to the environment than chemical 2, then all three evaluation parameters of chemical 1 can be higher than those of chemical 2. However, it is rare to have a large group of chemicals where all included parameters make it possible to make a linear ranking purely based on the 'unweighted' parameters, since one chemical could, for example, be more toxic but less persistent than its counterpart. This problem is traditionally solved by weighting toxicity in relation to persistence; otherwise it would not be possible to compare indistinguishable chemicals.

The HDT deals with indistinguishable chemicals in larger sets by dividing the set of chemicals to be compared into groups or branches (Fig. 3) of distinguishable chemicals. In these groups/branches, all evaluation parameters of one chemical are lower than the same parameters of the chemical ranking higher within the same group/branch. This typically results in more than one group/branch, and the determination of the absolute order of the individual chemicals between groups/branches is not possible directly, which is why HDT is a partial ordering method.

To determine the absolute order, it is necessary to compare the individual groups of indistinguishable chemicals using methods like linear extension. By linear extensions the order theoretical probability can be derived, so that a chemical gets a certain rank as presented in Lerche et al. (2003). Please refer to Brüggemann et al. (2001) for further details on the HDT theoretical frame.

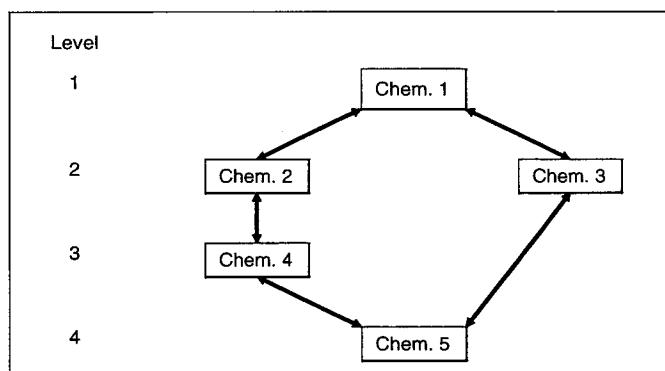


Fig. 3: Comparison of 5 chemicals. From the Hasse diagram (HD), it is seen that chemicals 3 and 4 rank above chemical 5. This means that all evaluation parameters on 3 and 4 are higher or lower than those of chemical 5. However, the absolute relation of the rank of chemical 3 in relation to chemical 4 is not possible to determine from the HD, since these are indistinguishable. Arrows indicate comparisons (distinguishable compounds)

2.7 CART

Existing examples of Classification And Regression Tree analysis (CART analysis) build on the assumption that the variation of properties follow a certain statistical distribution (as demonstrated in Bennet et al. 2000, Bennet et al. 2001, Eisenberg and McKone 1998). Using this assumption and Monte Carlo simulation, for example, it is possible to create a large set of hypothetical chemical compounds with properties varying according to a specified distribution and limits. As an alternative, large data sets for heterogenous chemicals could be analysed similarly.

The individual chemicals can be grouped according, for instance, to persistence or hazard. Using the set of properties and persistence, as one example, the CART analysis results in an evaluation tree created according to desired preferences (Fig. 4). The preferences can include minimizing false positives and virtually eliminating false negatives, or ensuring classification into a certain number of categories. As illustrated by the tree presented in Fig. 4, the persistence of a given chemical can be readily estimated with a well-defined certainty and in a structured manner from a subset of its properties.

3 Application of general principles for CRS to SMs

Davis et al. (1997) presented general principles, or a framework, of how to develop and select (for a specific purpose) chemical ranking and scoring systems. As far as we know, while other types of evaluation have been conducted, the only

CRS system that has been evaluated against the general principles is EURAM and this was done by the method developers themselves. They concluded that the EURAM method fulfils all the basic criteria, i.e. the 17 general principles (Hansen et al. 1999). These principles are reviewed below and applied to the SM framework and, as relevant, to the SMs and the other CRS methods included in this paper. More specific principles on exposure, human toxicity and ecotoxicity are dealt with in the background report (Larsen et al. 2002).

1) **There should be a clearly defined purpose.** For an SM, the purpose is to identify those emissions that are likely to contribute significantly to the impact category indicators for ecotoxicity and human toxicity prior to using an associated characterisation method with more data requirements.

2) **There should be compatibility with the risk assessment paradigm.** For an SM, it is relevant that it is compatible with the methods used and principles of LCIA for ecotoxicity and human toxicity indicators. Apart from the conceptual background and the purpose, these methods and principles are, to a large degree, currently inspired by, and to some degree compatible with, the risk assessment paradigm (e.g. Olsen et al. 2001).

3) **Uncertainty should be acknowledged and assessed.** Like dedicated CRS systems, SM results are highly uncertain, but the crucial point is not the absolute results, or scores, but the consistency with the associated characterisation method.

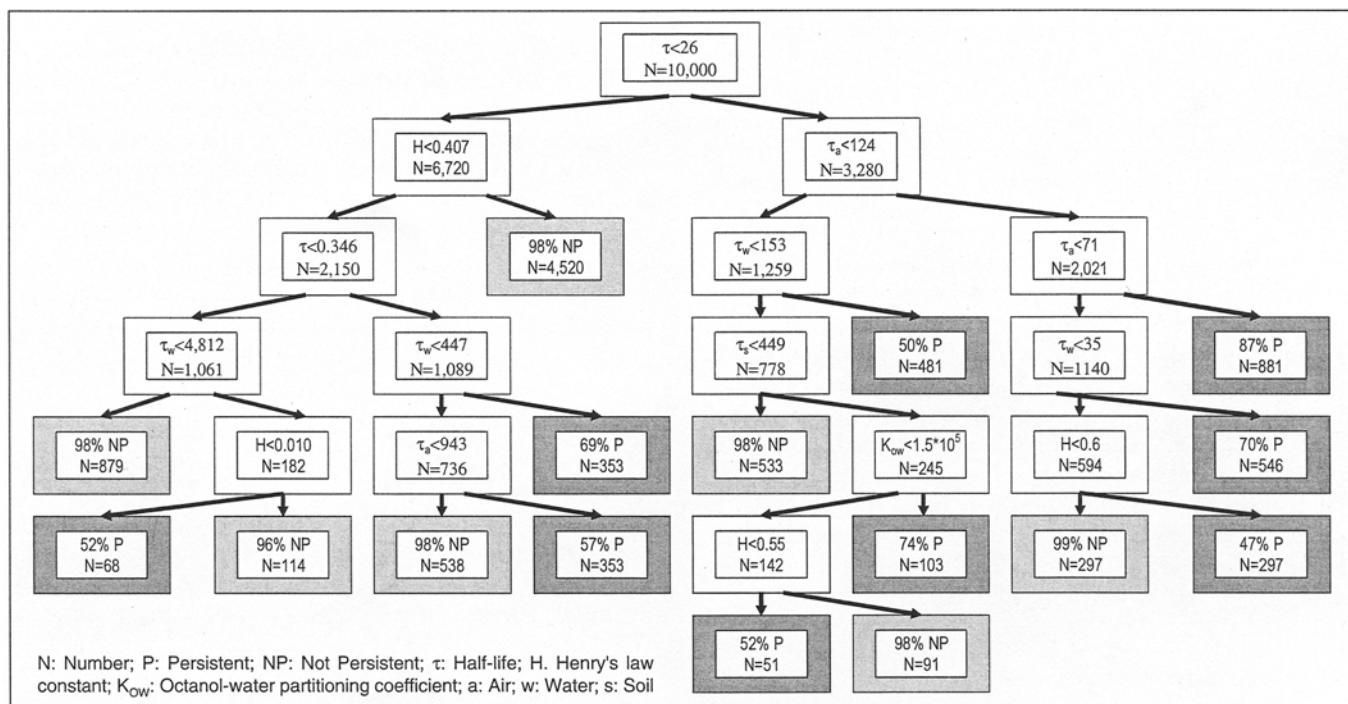


Fig. 4: Example of CART tree for emissions to air, from the mode of entry approach (from Bennett et al. 2001). The tree should be interpreted in that way that classification starts at the top, proceeding by evaluating the first inequality. If the inequality is true (i.e. answer is yes) the evaluation proceeds along the left branch. If the inequality is false (i.e. answer is no) the evaluation then proceeds along the right branch until reaching a terminal node (classification). Associated with each terminal node, dark grey for classification as persistent or light grey for classification as non-persistent, is a percentage indicating the confidence limit of the classification of the terminal node in question

4) The role of professional judgement should be acknowledged. Theoretically, the aim must be that professional judgement is only involved in the creation and evaluation of the SM (and possible updating) and that the user should be able to use the method without the need for specific scientific judgement. In practice, this is a rarely achievable objective in a pure sense.

5) There should be a broad consideration of effects. While LCIA addresses a very wide range of potential environmental impacts, the SMs are only supposed to cover the toxicity-related impact categories. Even for these, characterisation methods can operate with several subcategories, e.g. acute aquatic toxicity, chronic terrestrial toxicity and chronic aquatic toxicity, see de Koning et al. (2002), for more examples. It may therefore be necessary for an associated SM to deal with more than one 'scoring group' for each of the issues of ecotoxicity and human toxicity to be consistent with the characterisation method. But, it should be stressed that the aim is to keep the SM as simple as possible, i.e. not more sophisticated than just sufficient to be consistent with the characterisation method. Of the SMs dealt with in this paper, the EDIP-selection method deals with one 'scoring group' for each of the issues of ecotoxicity and human toxicity, whereas the Priofactor divides the human toxicity part in two 'sub-scoring groups', i.e. tox-priofactor (air) and tox-priofactor (soil/water). This sub-categorising is not directly based on effect type, but on exposure way. In the CPM-selection method, which only deals with ecotoxicity, two 'scoring groups' exist. Here, the differentiation is based on different behaviour in the environment (fate) for two different substance groups, i.e. organic and inorganic substances.

6) The role of valuation in aggregation and weighting should be recognised. Here, Davis et al. (1997) states that 'it generally is preferable that a chemical ranking system does not aggregate across major effect types', but in case that it should be done, the procedure should be transparent. The main argument is that aggregation and weighting require value judgement (political assessment) and when dealing with risk assessment, this judgement is dependent on the risk management context. For an SM, we assess that aggregation across ecotoxicity and human toxicity may not be a good idea, unless the results are correlated, because characterisation methods typically distinguish at least between these two and the selection would therefore most probably become inconsistent with the characterisation. The only CRS method included in this paper that includes the option to aggregate across ecotoxicity and human toxicity is the WMPT method. Here, it was done in a transparent way, although this was not retained in the context of the LCA-related publications, see section 2.5.

7) Methods and outputs should be transparent. It is preferable that an SM is based on transparent theoretical logic which is consistent with the characterisation method, transparent in mathematical formulations, and is well documented.

8) Method should be neutral to data availability. Davis et al. (1997) stated that a CRS method 'should not systematically 'punish' or 'reward' chemicals with extensive data versus

chemicals with no data'. This principle is in line with the aspiration of LCIA to provide a best estimate of the impact, while it deviates from the common procedure of risk screening where an approach is generally adopted that penalises substances with poor data availability. Generally, and including SMs, there are some mostly unsolved problems on how to deal with a lack of data, especially a lack of measured data. We assess that for SMs, one way to approach a solution to this problem is further use of the available state of the art QSARs to fill gaps, following hierarchy procedures (such as in the WMPT). In addition, as long as the number of false positives does not become unacceptably high, conservative handling of chemicals with no data is acceptable for an SM because it does to some degree help 'catch' possible false negatives.

9) Method should accommodate extreme variability in data availability across chemicals. Davis et al. (1997) pointed out that CRS 'systems must distinguish high-threat chemicals from chemicals with missing data'. This is also important for an SM to avoid bias in the ranking of emissions, as described above in principle 8. However, it should be noted that chemicals with a relatively low toxicity may contribute significantly to the toxicity-related impact potentials in an LCA, because the emitted amounts and the actual context (e.g. no highly toxic chemicals involved) can have a large influence on the outcome. Additionally, the impact potentials of the other impact categories included in the assessment may be low in an actual case study.

10) A tiered approach is practical and desirable. One could state that introducing an SM in an LCIA creates a two tiered approach, i.e. tier 1 comprises selection and tier 2 comprises characterisation. In cases where a lot of chemical emissions have to be ranked it may be practical, time saving and desirable to use a two tiered SM approach with increasing demand on e.g. data amount and quality. In cases where the number of emissions to be ranked is low, tier 1 might be skipped. Another possibility is to include a pre-selection in tier 1, e.g. that emissions of metals are pre-selected for characterisation based on the pre-existence of characterisation factors or high quality data sufficient to carry out a higher tier assessment. None of the existing SMs and the other CRS systems described in this paper are explicitly tiered (excluding CPM as a joint method).

11) Similar effects/exposure categories should be assessed across tiers. To avoid too many false positives and false negatives, the same main effects and exposure categories should be assessed in the SM (including eventually tiers) as in the associated characterisation method. To what degree this needs to be done for an SM depends on the complexity of the associated characterisation method and especially on the consistency, i.e. degree of agreement in prioritization of emissions between the SM and the characterisation method. All three SMs dealt with in this paper (EDIP-selection, Priofactor and CPM-selection) include effect and partly exposure (biodegradation and bioaccumulation), but only the Priofactor includes the emitted amount.

12) Critical information should be preserved. Critical information, e.g. indicators of data quality in scoring for each

emission and decisive data for each scoring, may become important when reviewing an LCA study.

13) Data selection guidelines should be specified. Data selection guidelines pointing out acceptable data sources, establishing a hierarchy of data sources, rules for manipulation, etc. are very important in order to avoid bias and errors in the ranking of the emissions. For SMs, at least some degree of consistency with the data sources used for the associated characterisation method is desirable, although the data demand of the SM should be lower in quantity and arguably also in quality. The SMs and the other CRS methods described in this paper include data selection guidelines to varying degrees. The EDIP-selection method points to a single source (i.e. the EU list of dangerous substances) as the main basis. The PrioFactor method is not that specific (especially not for ecotoxicity) but points out some quality demands for ecotoxicity data (preferably based on standard OECD tests) and some guidelines for the use of QSARs on BCF. The CPM-selection method is to a lesser degree specific on data selection guidelines than the two other SMs mentioned above. For the two non-selection methods dealt with here, it should be mentioned that the WMPT method is very detailed in defining data sources and a hierarchy of quality. The EURAM method uses IUCLID (IUCLID 2000) as an exclusive data source, given its policy support role and the relevance of this source. Data selection guidelines will probably have to be developed within OMNIITOX for a 'new' SM, and here, the work done by Pedersen et al. (1995) on data selection strategy and quality (only ecotoxicity), and by Larsen et al. (1999a) (both ecotoxicity and human toxicity), and the WMPT approach, could be a good starting point/basis.

14) Method should be theory driven as well as data driven. According to Davis et al. (1997), a CRS method should as the starting point be based on a theoretical logic, which afterwards is meshed with data availability considerations. An SM should be based on a theoretical logic in accordance with the associated characterisation method to the degree necessary to give the required consistency between the ranking of the two methods. But it must be stressed that the consistency on prioritization of emissions and the need for a low data demand of the SM (as compared to the characterisation method) are more important criteria.

15) Sensitivity analysis should be performed. This point is aimed at development of CRS systems in general and is therefore also relevant when developing an SM. Sensitivity analysis may have the strongest priority when dealing with comparison between an SM and its associated characterisation method, e.g. looking for key parameters ('drivers') to minimise this to essential data and trying to find explanations for differences in prioritization of emissions.

16) Pre-selection of chemicals should be consistent with the CRS method. The pre-selection of emissions to be included in the ranking and scoring procedure may result in false negatives if this pre-selection is not well founded. Pre-selection is not included in any of the SMs described here, but, as implicitly in many LCIA applications, the EDIP-selection

method includes some kind of 'pre-exclusion' mainly based on common sense (well-founded and based on long experience), e.g. water and CaCO_3 are not taken into account as may also be the case implicitly in other CRS methods.

17) The impact of scaling should be considered. Whether the scale used in an SM is ordinal (e.g. 1, 2, 3 ...), nominal (e.g. 'yes' or 'no'), interval (e.g. Celsius scale of temperature) or ratio (having true zero point, e.g. Kelvin scale of temperature) is not that important for an SM, as long as the SM is consistent with the associated characterisation method. However, the scaling used and mathematics of combining values to calculate a score may have significant influence on the data demand and the results.

4 Developed Criteria for the Evaluation of SMs

To support the evaluation of the different methods, a set of criteria for performance of an SM is proposed from the above principles. A good SM is one that is:

- Consistent with the associated characterisation method in prioritization/ranking, methodology and substance data requirement
- Transparent, i.e. well-documented and manual calculation is possible
- Applicable to different chemical groups
- Operational with a data requirement that
 - is modest in demand on scope and quality of the substance data
 - comprises data also needed for calculation in characterisation step
 - is focused on data of high availability and with possible prediction, e.g. applying QSARs
- User friendly
 - high practicability – easy in use, preferably with end-user software existing
 - lower time demand as compared to the characterisation method
 - lower requirements to specific scientific skills of the user

A quantitative internal weighting of the criteria is of little relevance in the OMNIITOX project, since the number of existing SMs is very low and it is a stated goal of the OMNIITOX project to develop a new stand-alone SM or an SBM (maybe as a tiered approach) used as a kind of combined SM and simple characterisation method in support of the BM. There is thus no need for identification of the best among existing methods. A ranking of the criteria provides, however, useful guidance for the further development of SMs or simple characterisation methods used to support more advanced ones. The following ranking is recommended:

1. Validity of the method
 - a) appropriately prioritizes the chemical emissions which are evaluated as significant in the associated characterisation step (no false negatives)
 - b) low in false positives
2. Applicable to different chemical groups, lower in data demand than the characterisation method, and based on data of high availability
3. User friendly and transparent
 - a) lower in time demand than the characterisation and easy to use
 - b) skilled scientific background not needed, as far as appropriate

5 Evaluation of the Chosen Methods

The three SMs and the other CRS methods are evaluated against the developed criteria. The result of this evaluation is shown in Table 7.

Table 7: Results of the evaluation of the SMs and the two other CRS methods

Criteria	EDIP-selection	Priofactor	CPM-selection	EURAM	WMPT
1a. Consistency in prioritization (validity)	+	+++	n.a.	n.r.	n.r.
2a. Applicable to different chemical groups	++++	+	+++	+	++
2b. High data availability	+++	++	++	+	+++
2c. Low data demand	++++	++	++	+	+
2d. Data useable in characterisation	+	+++	+++	n.r.	n.r.
3a. Low in time demand and easy to use	+++	++	++	+	(++)
3b. Skilled scientific background not needed	+++	++	++	+	+

++++: Very high degree of fulfilment
+++: High degree of fulfilment
++: Moderate degree of fulfilment

+: Low degree of fulfilment
n.a.: Not assessed
n.r.: Not relevant

6 Discussion

The Hasse diagram technique and the CART analysis have not been assessed using the developed criteria and are therefore not included in Table 7. This is because these approaches represent frameworks rather than operational methods in an LCA context. Parameters used in these approaches are not predefined and this makes it impossible to apply most of the criteria. It is however possible to assess transparency and user friendliness to some degree, as both methods are well documented and other practical examples do exist. The underlying methodologies and available studies suggest that these approaches may help reduce the data requirements and type compared to the needs of the characterisation methods. Once the methods have been applied to develop the SMs, these will typically be straightforward to use. Manual use may be undesirable, but the methods could be implemented in straightforward spreadsheets.

As shown in Table 7, none of the evaluated methods have a very high or even high fulfilment of all criteria. All have weaknesses and strengths. Here we focus on the parts of the different methods that are strong in the sense of performance as an SM. A more detailed analysis of pros and cons can be found in Larsen et al. (2002).

The EDIP-selection method is in principle applicable to all chemical groups as based on R-phrases. This basis also gives rise to high data availability (very high if QSAR calculated data are included – especially for the ecotoxicity part). The methodology is based on a very simple scoring system and it is therefore easy to use, low in time demand and specific scientific skilled background can be considered minimal.

The Priofactor method has a high consistency with the associated characterisation method. In the evaluation example quoted in this paper, it only comes up with a few false positives and a few false negatives. The false negatives only account for below 1% (ecotoxicity) or below 4% (human toxicity) of the total impact, as calculated by the associated characterisation method. The main reason for the high consistency is probably that this SM, to a certain degree, is based on the same methodology and principles as the characterisation method, but makes use of acute data instead of chronic. Some of the (perhaps less significant) key parameters are excluded and the emitted amount is included. The data needed for the Priofactor therefore has a high usability with the associated characterisation method, i.e. the EDIP method.

The CPM-selection method distinguishes between organics and inorganics (primarily metals) and includes special parameters

for inorganic chemicals, like covalent bonding and redox sensitivity, which is not seen in any of the other methods described here. This method therefore has a relatively high applicability to different chemical groups. As is the case in all the other methods, amphiphilic and dissociating chemicals are not specifically addressed, i.e. they are treated like any other organic chemical. Similar to the Priofactor, the CPM-selection method uses the same methodology/principles as the associated CPM-characterisation method, but in a more simple way, i.e. acute toxicity data instead of chronic, etc.

For the two evaluated non-SMs, the main feature that distinguishes them from the three evaluated SMs is the way in which the fate part is handled. Both of them apply multimedia fate models, i.e. Mackay level I for EURAM and Mackay level III for WMPT, which have a higher data demand (especially level III) than the more simple key parameter based approaches used in the SMs. However, this may not prove to be a hindrance if QSARs and simple data guidelines are used.

In general, it can be stated that the methods evaluated here are semi-quantitative apart from the Priofactor where the ecotoxicity part must be considered as almost fully quantitative. Only the Priofactor and the EURAM method take the emitted amount into account, and only the CPM-selection method takes the special chemistry of the inorganic compounds into considerations. However, how to account for emitted quantity using the WMPT and similar methods in an LCA context, and also for addressing substances such as metals independently have been demonstrated, see section 2.5.

7 Conclusion, Recommendations and Outlook

In conclusion, it is recommended that the principles on how to develop and to select chemical ranking and scoring systems should be considered when developing the approach for selection (simplified methods, SMs) in LCIA of toxic emissions. In addition, a number of more specific recommendations are presented for developing new SMs:

- emitted amount, as quantified in the inventory results of a specific LCA-study, should be taken into account
- very high data availability for the underlying substance data is a prerequisite (based on e.g. R-phrases, K_{ow} and/or QSARs)
- method should be low in false positives and create no false negatives, or false negatives should only account for an insignificant part of the impact potential as calculated by the associated characterisation method

It is recommended that further work on selection includes a comparison with the use of the characterisation method for selection based on factors calculated with an extensive use of QSAR for estimation of the substance input parameters. Other types of selection approaches should be tested in parallel and, therefore, the following methods (or similar ones) are recommended to be included in the further work:

- EDIP-selection
- PrioFactor
- CPM-selection
- EURAM
- WMPT
- Hasse diagram technique
- CART?

The Hasse diagram technique may be used as a 'baseline' CRS method (excluding value choices or subjective rules) when comparing the ranking of the different methods. The CART analysis has only been demonstrated in the context of persistence (fate), but a statistical approach may be beneficial for an SM associated with a coming newly developed OMNIITOX characterisation method (BM) and CART could become relevant in this context.

Experience from application of the few existing SMs is very limited. Of the approaches classified as SMs, only the EDIP-selection method and the PrioFactor have been evaluated against the associated characterisation method and only on a single case inventory. In a broader sense, some Chemical Ranking and Scoring (CRS) approaches have been tested against more sophisticated approaches and related insights could prove beneficial.

On the basis of the developed criteria for performance of SMs and the recommended issues to be taken into account when developing new SMs, it is recommended that the research on selection focuses on the following issues:

- Is it possible to use simple data with very high availability (e.g. R-phrases, QSAR calculations) and simultaneously achieve high consistency with the associated characterisation method?
- What are the key parameters for a stand alone SM (e.g. persistence, bioaccumulation, toxicity and emitted amount) to include in an SM and how? Can it be done without subjective/expert choices (e.g. Hasse diagram technique) or are such choices essential?
- Is explicit multimedia fate modelling necessary or, as generally considered, is a simple key parameter approach sufficient to achieve high consistency with the associated characterisation method?
- Will a statistical approach like CART analysis, linear regression or other approaches like mega variate data analysis be appropriate to achieve high consistency with the associated characterisation method?

A framework for developing and evaluating SMs has been presented together with methods which are intended to act as selection methods. No conclusive recommendations are given on which route to follow to derive an SM, but it remains clear that, in order to facilitate inclusion of the toxic-

ity-related impact categories (i.e. ecotoxicity and human toxicity) in LCA studies on a level similar to the well implemented impact categories (e.g. global warming), there is a need for further research on, and experience with use of, selection methods.

Acknowledgement. The authors wish to thank the anonymous reviewer for a well structured and relevant review. The authors also gratefully acknowledge the financial support of the European commission through the 'Sustainable and Competitive Growth'-research programme given to the OMNIITOX project and the financial support from the Danish Ministry of Science Technology and Innovation given to KEMI (Centre for chemicals in industrial production).

Glossary

BCF	BioConcentration Factor
BDF	BioDegradation Factor
BM	Base Model
CART	Classification And Regression Tree analysis
CPM	Competence Centre for Environmental Assessment of Products and Materials Systems
CRS	Chemical Ranking and Scoring
EC ₅₀	Effect Concentration (50% of test organism affected)
EDIP	Environmental Design of Industrial Products
EPA	(US) Environmental Protection Agency
EU	European Union
EURAM	EU Risk Ranking Method
HDT	Hasse Diagram Technique
K _d	Soil adsorption coefficient
LC ₅₀	Lethal Concentration (50% of test organism dead)
LCA	Life Cycle Assessment
LCIA	Life Cycle Impact Assessment
LogK _{ow}	The logarithm of the Octanol/Water partition coefficient
N.C.	Not Classified
PBT	Persistence, Bioaccumulation and Toxicity
pKa	Dissociation constant for acids (and bases)
QSAR	Quantitative Structure Activity Relationship
R-phrases	Risk phrases
SBM	Simple Base Model
SM	Selection Method
WMPT	Waste Minimization Prioritization Tool

References

Bennett DH, McKone TE, Kastenberg W (2001): CART Screening level analysis of persistence: A case study. In: Lipnick R L, Jansson B, Mackay D, Petreas M (editors) (2001): Persistent, bioaccumulative and toxic chemicals II: Assessment and new chemicals. ACS symposium series 773. Published by American Chemical Society, Washington, DC. ISBN 0-8412-3675-5

Bennett DH, McKone TE, Kastenberg W (2000): Evaluating multimedia persistence: Classification and regression tree analysis. Environ Toxicol Chem 19, 810-819

Brüggemann R, Bücherl C, Pudenz S, Steinberg CEW (1999): Application of the concept of partial order on comparative evaluation of environmental chemicals. Acta hydroch hydrob 27 (3) 170-178

Brüggemann R, Halfon E, Welzl G, Voigt K, Steinberg CEW (2001): Applying the concept of partial ordered sets on the ranking of

near-shore sediments by a battery of tests. *J Chem Inf Comp Sci* 41 (4) 918–925

Bunke D, Gensch C-O, Möller M, Rüdenauer I, Ebinger F, Graulich K (2003): Assessment of Toxicological Risk due to Hazardous Substances: Scoring of Risk Phrases. *Int J LCA* 8 (1) 6–7

Davis G, Fort D, Hansen BG, Irwin F, Jones B, Jones S, Socha A, Wilson R, Haaf B, Gray G, Hoffman B (1997): Framework for Chemical Ranking and Scoring Systems. Chapter 1 in: Chemical ranking and scoring: Guidelines for relative assessment of chemicals, Swanson M B and Socha A C (eds.), pp. 1–29. SETAC-US. Proceedings from the Pellston workshop on Chemical ranking and scoring, 12–16 February, 1995. Sandestin, Florida

de Koning A, Guinée J, Pennington DW, Slesewijk A, Hauschild M, Molander S, Nyström B, Pant R, Schowanek D (2002): Methods and typology report Part A: Inventory and classification of LCA characterisation methods for assessing toxic releases. Contribution to Work-package 7 of the OMNIITOX Project as part A of appropriate deliverable D11

EC (1996): Environmental Risk Assessment. Technical Guidance Document in support of Commission Directive 93/67/EEC on Risk Assessment for New Notified Substances and Commission Regulation (EC) No. 1488/94 on Risk Assessment for Existing Substances (ed EC), pp 241–504. Office for Official Publications of the European Communities, Luxembourg

EC (2001): Adapting to technical progress for the 28th time Council Directive 67/548/EEC on the approximation of laws, regulation and administrative provisions relating to the classification, packaging and labelling of dangerous substances (Commission Directive 2001/59/EC, August 6)

EEC (1967): Approximation of laws, regulations and administrative provision relating to the classification, packaging and labelling of dangerous substances (Council Directive 1967/548/EEC, June 27)

Eisenberg JN, McKone TE (1998): Decision tree method for the classification of chemical pollutants: Incorporation of across-chemical variability and within-chemical uncertainty. *Environ Sci Technol* 32 (21) 3396–3404

Eriksson Å (1999): A comparative study of four models for eco-toxicological impact assessment in LCA – A case study of tyre emissions. Honour thesis in environmental science. Tutor Sverker Molander. June 1999. Göteborgs Universitet, Sweden

Fomsgaard I, Sørensen PB (1999): The Hasse diagram. In: Comparing Environmental Risk indicators for pesticides. Results of the CAPER project. Centre for Agriculture & Environment, The Netherlands. Report CLM 426-1999. ISBN 90-5634-106-5

Guinée JB, de Koning A, Pennington DW, Rosenbaum R, Hauschild M, Olsen SI, Molander S, Bachmann T M, Pant R (2004): Bringing Science and Pragmatism together in a Tiered Approach for Modelling Toxicological Impacts in LCA. *Int J LCA* 9 (5) 320–326

Halfon E, Galassi S, Brüggemann R, Provini A (1996): Selection of priority properties to assess environmental hazard of pesticides. *Chemosphere* 33 (8) 1543–1562

Hansen BG, Van Haelst AG, Van Leeuwen K, Van der Zandt P (1999): Priority setting for existing chemicals: European Union risk ranking method. *Environ Toxicol Chem* 18, 772–779

IUCLID (2000): International Uniform Chemical Information Database. Public data on high volume chemicals. (existing chemicals). Year 2000 ed. European Chemicals Bureau. Ispra, Italy

Larsen HF, Birkved M, Hauschild M, Pennington D, Guinée JB (2002): Inventory of LCIA selection methods for assessing toxic releases. Methods and typology report part B. Contribution to Work-package 7 of the OMNIITOX Project as part B of appropriate deliverable D11. EC Project contract G1RD-CT-2001-00501 <<http://www.omniitox.net/>>

Larsen HF, Christensen FM, Rasmussen JO, Helweg C, Olsen SI (1999b): Screening tool for prioritization of chemicals to be included in a LCA study. Presented at the 9th Annual Meeting of SETAC-Europe, 25–29 May 1999 in Leipzig, Germany. Abstract No. 4c/001

Larsen HF, Christensen FM, Rasmussen JO, Helweg C, Olsen SI (1999a): CETOX-LCA-Handbook, LCA on 'chemical-heavy' products. CETOX, Hørsholm, Denmark

Lerche D, Sørensen PB, Larsen HS, Carlsen L, Nielsen O-J (2002): Comparison of the Combined Monitoring-based and Modeling-based Priority Setting Scheme with Partial Order Theory and Random Linear Extension for Ranking of Chemical Substances. *Chemosphere* 49 (6) 637–649

Lerche D, Sørensen PB, Brüggemann R (2003): Improved estimation of the ranking probabilities in partial orders using random linear extensions by approximation of the mutual ranking probability. *J Chem Inf Comp Sci* 43 (5) 1471–1480

Molander S, Lidholm P, Schowanek D, Recasens M, Fullana P, Christensen FM, Guinée JB, Hauschild M, Jolliet O, Carlson R, Pennington DW, Bachmann TM (2004): OMNIITOX – Operational life-cycle impact assessment models and information tools for life-cycle impact assessment practitioners. *Int J LCA* 9 (5) 282–288

Olsen SI, Christensen FM, Hauschild M, Pedersen F, Larsen HF, Tørsøv J (2001): Life cycle impact assessment and risk assessment – A methodological comparison. *Environmental Impact Assessment Review* 21, 385–404

Pedersen F, Tyle H, Niemelä JR, Guttmann B, Lander L, Wedebrand A (1995): Environmental Hazard Classification – Data collection and interpretation guide (2nd edition). TemaNord 1995:581. Nordic Council of Ministers

Pennington DW, Bare JC (2001). Comparison of chemical screening and ranking approaches: The Waste Minimization Prioritization Tool (WMPT) vs. Toxic Equivalency Potentials (TEPs). *Risk Analysis* 21 (5) 897–913

Pennington DW, Bare JC, Knodell R, Carroll G, Martin T (2003): Evaluating Pollution Prevention Progress (P2P). Clean Technologies and Environmental Policy, in press

Sørensen PB, Mogensen BB, Gyldenkærne S, Rasmussen AG (1998). Pesticide leaching assessment method for ranking both single substance and scenarios of multiple substance use. *Chemosphere* 36 (10) 2251–2276

US EPA (1997) Notice of availability of waste minimization software and documents. EPA530-C-97-003, Federal Register, 62:33868

US EPA (1998) Notice of availability of draft RCRA waste minimization PBT chemical list. 530-Z-98-009, Federal Register, 63:216

van de Meent D, Struijs J, Sijm D (2002): Simple recipe for calculating multi-substance toxic pressure as indicator of ecological impact of toxic substances. Presented at the 12th Annual Meeting of SETAC-Europe, 12–16 May 2002 in Vienna, Austria. Abstract No. 65-07

Wenzel H, Hauschild M, Alting L (1997): Environmental Assessment of Products, First edn. Chapman & Hall

Received: October 9th, 2003

Accepted: August 15th, 2004

OnlineFirst: August 16th, 2004